6-1990

Characterization and expression of the cbbE' gene of Coxiella burnetii

Michael F. Minnick

Robert A. Heinzin

Marvin Eugene Frazier

Louis P. Mallavia

Characterization and expression of the *cbbE'* gene of *Coxiella burnetii*

MICHAEL F. MINNICK,[1] ROBERT A. HEINZEN,[1,2] MARVIN E. FRAZIER[2] and LOUIS P. MALLAVIA[1]*

[1] *Department of Microbiology, Washington State University, Pullman, WA 99164-4233, USA*
[2] *Battelle, Pacific Northwest Laboratories, Richland, WA 99352, USA*

(*Received 31 October 1989; revised 8 January 1990; accepted 6 February 1990*)

A gene which is unique to the QpRS plasmid from chronic isolates of *Coxiella burnetii* was cloned, sequenced, and expressed in *Escherichia coli*. This gene, termed *cbbE'*, codes for a putative surface protein of approximately 55 kDa, termed the E' protein. The *cbbE'* gene is 1485 bp in length, and is preceded by predicted promoter regulatory sequences of TTTAAT (−35), TATAAT (−10), and a Shine-Dalgarno sequence of GGAGAGA, all of which closely resemble those of *E. coli* and other rickettsiae. The open reading frame (ORF) of *cbbE'* ends with a UAA codon followed by a second in-frame UAG stop codon and a region of dyad symmetry which may act as a rho-factor-independent terminator. The ORF of *cbbE'* is capable of coding for a polypeptide of 495 amino acids with a predicted molecular mass of 55893 Da. The E' protein has a predicted pI of ~8·7, and contains a distinct hydrophobic region of 12 amino acid residues. *In vitro* transcription/translation and *E. coli* expression of recombinant plasmids containing *cbbE'* produce a protein of approximately 55 kDa. The *in vivo* expression of *cbbE'* yields a novel protein that can be detected on immunoblots developed with rabbit antiserum generated against purified outer membrane from *C. burnetii*. DNA hybridization analysis shows that *cbbE'* is unique to the QpRS plasmid found in chronic isolates of *C. burnetii*, and is absent in chromosomal DNA and plasmids (QpH1, QpDG) from other isolates of *C. burnetii*. A search of various DNA and amino acid sequence data bases revealed no homologies to *cbbE'*.

Introduction

The rickettsia *Coxiella burnetii* is an obligate intracellular parasite that causes Q fever in humans. Although Q fever generally manifests as an acute, self-limiting, flu-like illness, approximately 5% of the reported cases develop chronic forms of the disease that may culminate in endocarditis and/or hepatitis (Kimbrough *et al.*, 1979). Isolates from acute cases of Q fever are identical to the Nine Mile phase I isolate with respect to their lipopolysaccharide (LPS) (Hackstadt, 1986) and by having a cryptic plasmid, termed QpH1 (Samuel *et al.*, 1983).

In contrast, isolates from chronic cases have LPS that is electrophoretically and antigenically distinct from that of acute isolates (Hackstadt, 1986; Moos & Hackstadt, 1987). Chronic isolates also lack the QpH1 plasmid; instead, they have a unique plasmid termed QpRS, or have chromosomal DNA sequences with homology to both QpH1 and QpRS (Samuel *et al.*, 1985). The homology of the chromosome of the plasmidless-chronic isolates with QpH1 and QpRS suggests that the plasmids may have integrated into the chromosome of these isolates (E. Savinelli & L. P. Mallavia, unpublished results). The QpH1 and QpRS plasmids are approximately 36 kb and 39 kb, respectively, in size, and possess unique and common DNA sequences as determined by Southern blot analysis (Samuel *et al.*, 1985).

A third plasmid, termed QpDG, has recently been purified from *C. burnetii* isolates from feral rodents collected near Dugway, Utah, USA (L. R. Hendrix, J. E. Samuel & L. P. Mallavia, unpublished results). Isolates that possess the QpDG plasmid do not cause apparent disease in guinea-pigs and have not yet been isolated from humans.

These observations suggest a correlation between plasmid type and the chronic or acute nature of disease per given isolate (Samuel *et al.*, 1985; Vodkin *et al.*, 1986). Our hypothesis is that DNA sequences which are unique to each plasmid type may code for virulence determinants which in turn determine the acute or chronic nature of disease. To investigate this possibility,

Abbreviations: IVTT, *in vitro* transcription/translation; OM, outer membrane; ORF, open reading frame; SD, Shine–Dalgarno.

The nucleotide sequence data reported in this paper have been submitted to GenBank and have been assigned the accession number M29982.

DNA sequences that are specific to either the QpH1 or QpRS plasmids are being characterized.

Only two genes, *htpA* and *htpB*, have been cloned and sequenced from the chromosome of *C. burnetii*. These genes were found to comprise a heat-shock operon which codes for both a 62 kDa and a 14 kDa protein (Vodkin & Williams, 1988). In this paper the first expression and sequence analysis of a *C. burnetii* plasmid gene is presented. The gene is termed *cbbE'*; it is unique to the QpRS plasmid of chronic isolates and codes for a putative surface protein of approximately 55 kDa.

Methods

Bacterial strains, plasmids and growth conditions. Propagation and harvesting of *C. burnetii* was done as previously described (Samuel *et al.*, 1983). *C. burnetii* plasmids and chromosomal DNA were isolated and purified as previously reported (Samuel *et al.*, 1983) from acute or chronic isolates, including phase I Nine Mile RSA-493(QpH1), Priscilla Q177(QpRS), Dugway 7E9-12(QpDG), and the plasmidless-chronic isolates Ko Q229, S Q217, G Q212 and L Q216. *Escherichia coli* strain DH5α or DH5αF' (BRL) were used as hosts for all recombinant DNA. *E. coli* cultures containing plasmids were grown at 37 °C in Luria–Bertani (LB) medium under antibiotic selection with ampicillin (100 μg ml^{-1}) for the pUC vectors (Vieira & Messing, 1983) and kanamycin (25 μg ml^{-1}) for the cosmid vector pHK17 (Klee *et al.*, 1983). *E. coli* strain SG932 (*lon-100*) (Goff *et al.*, 1984) was used for *in vivo* expression of the cloned *cbbE'* gene, and was grown at 30 °C in LB with ampicillin (100 μg ml^{-1}).

Construction of recombinant DNA. *Eco*RI (BRL) fragments of QpRS were separated by electrophoresis through 0·9% (w/v) agarose gels prepared by standard methods (Maniatis *et al.*, 1982), excised from agarose gels and purified by Geneclean (Bio 101). Purified DNA was then cloned by standard procedures into pUC9 or pUC19 (Maniatis *et al.*, 1982). *E. coli* DH5α was transformed with the ligation mixture following $CaCl_2$ treatment (Davis *et al.*, 1980). Colonies that contained pUC recombinant plasmids were detected on LB with ampicillin (100 μg ml^{-1}), isopropyl β-D-thiogalactopyranoside (IPTG; 0·3 mM) and Bluo-gal (BRL). All plasmid DNA was isolated from *E. coli* DH5α by alkaline extraction (Birnboim & Doly, 1979). Plasmid DNA was separated from chromosomal DNA by two CsCl density-gradient centrifugations.

DNA hybridization. Southern blot analysis was used to detect the *cbbE'* gene in chromosomal and/or plasmid DNA isolated from typical isolates of the four known strains of *C. burnetii*. The QpH1 and QpRS plasmids, cloned in their entirety into the *Sal*I site of the cosmid vector pHK17 (Klee *et al.*, 1983) to produce pQH1 and pQRS, respectively, were also probed. The internal 695 bp *Pst*I fragment of the *cbbE'* gene (Fig. 1) was isolated from the pQME1 subclone as described above, and then labelled with [α-^{32}P]dCTP by a random primer extension kit (New England Nuclear). The DNA to be probed was digested to completion with restriction enzymes, electrophoresed in 1% agarose, and then denatured and transferred to nitrocellulose (0·45 μm pore size, Schleicher & Schuell) by the method of Southern (1975). The nitrocellulose was prehybridized for 1 h at 68 °C, then hybridized for 16 h at 68 °C with approximately 10^6 d.p.m. of the *Pst*I fragment probe ml^{-1} as previously described (Samuel *et al.*, 1983). The filter was then washed four times at high stringency, for 30 min at 68 °C with 45 mM-sodium chloride plus 4·5 mM-sodium citrate (0·3 × SSC) containing 0·1% SDS and 5 mM-EDTA.

Gene expression and immunoblot detection. A prokaryote-directed *in vitro* transcription/translation (IVTT) system was used as directed by the manufacturer (Amersham) to analyse proteins coded on the intact ε' fragment and its various subfragments. Translated proteins were labelled with 30 μCi (1·11 MBq) [^{35}S]methionine (New England Nuclear) for 1 h at 37 °C. IVTT samples of approximately 2×10^5 d.p.m. were boiled for 5 min in an equal volume of electrophoresis sample buffer, then separated by SDS-PAGE (12·5%, w/v, acrylamide) (Laemmli, 1970). The ^{35}S-labelled proteins were visualised by autoradiography following overnight exposure to X-omat X-ray film (Kodak).

In vivo expression of the E' protein was achieved in *E. coli* SG932, which is defective in the degradation of abnormal proteins (Goff *et al.*, 1984). A 250 μl volume of a fresh overnight culture of SG932 containing pQME3 served as the inoculum for 5 ml LB medium containing 1 mM-IPTG to induce expression. Mid-exponential-phase cells (OD_{600} ~0·6) were harvested by centrifugation and washed twice with cold 0·15 M-NaCl, then stored at −20 °C until needed. Following protein assay by the Lowry method, the pellets were resuspended in SDS-PAGE sample buffer, boiled for 5 min, then electrophoresed in SDS-PAGE gels (12·5% acrylamide) containing 30 μg protein per lane (Laemmli, 1970). Protein profiles were then analysed for expression of the E' protein by immunoblot analysis (Towbin *et al.*, 1979). Unfixed and unstained gels were electrotransferred to prewetted nitrocellulose (0·45 μm pore size) for 16 h at 200 mA, with constant cooling. The filters were dried, then immersed for 1 h in phosphate-buffered saline (PBS) with 0·3% (v/v) Tween 20 and 2% (w/v) nonfat milk. The filters were then exposed for 16 h at 25 °C to rabbit antiserum against Priscilla outer membrane, diluted 1:1000 in PBS. Filters were washed five times in PBS, then exposed for 1 h to horseradish-peroxidase-conjugated goat anti-rabbit immunoglobulin (Tago) diluted 1:2000 in PBS. After four washes in PBS, the filters were developed in 50 ml PBS containing *o*-dianisidine (0·5 mg ml^{-1}) plus 75 μl 30% H_2O_2.

DNA sequence determination and analysis. The nucleotide sequence of both DNA strands of *cbbE'* was determined by the dideoxy chain-termination method of Sanger *et al.* (1977) using [α-^{32}P]dCTP (3000 Ci $mmol^{-1}$; 111 Tbq $mmol^{-1}$) as the radiolabel, and Sequenase (United States Biochemical). Single-stranded pGEM7 templates containing regions of the ε' fragment were primed using a T7 promoter primer (Promega). Double-stranded pGEM7 or pUC19 recombinants containing ε' subfragments were primed using pUC forward or reverse primers (Promega). All double-stranded templates were purified twice on CsCl gradients. When necessary, oligonucleotide sequencing primers (17–25-mers) were synthesized using an Applied Biosystems DNA synthesizer (model 380A). Sequence data were compiled and analysed using CAGE/GEM (Battelle, Pacific Northwest Laboratories, Richland, Washington, USA), the University of Wisconsin Genetics Computer Group (Deveroux *et al.*, 1984) and Pustell (International Biotechnologies) DNA sequence analysis programs.

Antibody preparation against C. burnetii outer membrane (OM). An adult male New Zealand White rabbit was used to generate polyclonal antiserum against isolated outer membrane (OM) of the Priscilla isolate of *C. burnetii*. The OM was purified by modification of an SDS extraction procedure (Hurlbert & Gross, 1983), in which approximately 30 mg (dry weight) of freshly isolated cells was rinsed in 20 vols of 10 mM-HEPES buffer with 10% (w/v) sucrose (pH 7·4 at 4 °C), then pelleted by centrifugation at 12100 *g* for 45 min at 4 °C. The pellet was resuspended in 2 ml 10 mM-HEPES and extracted in 0·4% (w/v) SDS. To liberate OM, four sonications with a Sonifier 250 (Branson) were done on ice for 1 min at constant cycle with a 1 min cooling interval. Cell debris was removed from the suspended membranes by centrifugation at 12100 *g*, as described above. The supernatant was then ultracentrifuged at 260000 *g* for 1 h at 4 °C, in a 60Ti rotor (Beckman). The insoluble transparent pellet was washed twice in the

10 mM-HEPES buffer. The final pellet was resuspended in physiological saline (0·15-M-NaCl). Approximately 200 μg of OM protein in saline, as determined by the Lowry method, was emulsified in an equal volume of incomplete Freund's adjuvant and delivered as three subcutaneous 1 ml injections above the shoulders of the rabbit. Two subsequent injections of the same concentration and volume were administered in the same manner at 2-week intervals. Antiserum was collected 2 weeks after the last injection.

Results

Cloning of the cbbE′ gene

A 3·6 kb *Eco*RI fragment of QpRS (the fifth-largest *Eco*RI fragment, hence the designation ε′), was cloned in both orientations in pUC19, to produce pQME1 and pQME2 (Fig. 1). The cloned ε′ fragment was mapped by restriction enzymes (BRL), and various subfragments were cloned into pUC19 for IVTT analysis. A 3·2 kb *Eco*RI–*Hinc*II subfragment was directionally cloned into pUC9 to form pQME3 (Fig. 1) for *in vivo* expression of *cbbE′*.

Southern blot analysis

The ε′ fragment was previously shown to contain unique sequences when probed with α-^{32}P-labelled QpH1 in Southern blot analysis (Samuel *et al.*, 1985). Further analysis showed that an internal *Pst*I subfragment of ε′ was unique to the QpRS plasmid by DNA hybridization (Fig. 2). Under the stringency employed, homologous sequences were detected in the DNA restriction digests of pQME1, of the QpRS-cosmid recombinant pQRS, and of the chromosomal DNA of the Priscilla isolate (Fig. 2). The latter sample probably contained nicked QpRS plasmid contaminant from the CsCl separations. Neither chromosomal nor plasmid DNA from either the Nine Mile isolate or the Dugway isolate had sequences homologous to the *Pst*I fragment, nor did the vector controls (Fig. 2). No homologous sequences were detected in the chromosomal DNA of the plasmidless-chronic isolates Ko, S, G or L, although a number of sequences with homology to QpRS were previously detected in these isolates by similar analyses (E. Savinelli & L. P. Mallavia, unpublished results).

Expression of the cbbE′ gene

Because of its uniqueness to QpRS, the ε′ fragment was characterized by IVTT. IVTT of pQME3 produced a protein of approximately 55 kDa that was specific to the rickettsial DNA insert, when analysed by SDS-PAGE (Fig. 3). A protein of the same molecular mass was produced in IVTT analyses of pQME1 and pQME2 (M. F. Minnick, R. A. Heinzen, L. P. Mallavia & M. E.

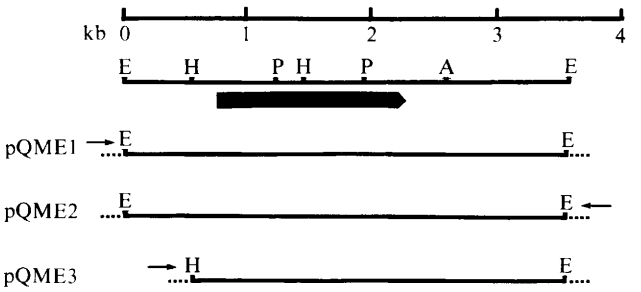

Fig. 1. Cloned DNA of the *C. burnetii* QpRS plasmid used to characterize and express the *cbbE′* gene. The rickettsial DNA insert is shown as a solid line with restriction sites indicated (A, *Acc*I; E, *Eco*RI; H, *Hinc*II; P, *Pst*I). The large arrow designates the location and direction of the *cbbE′* open reading frame (ORF). Broken lines represent the pUC19 vectors in plasmids pQME1 and pQME2, and represent pUC9 for plasmid pQME3. The direction of the *lacZ′* promoter with respect to the insert is indicated by a small arrow. Only the flanking restriction sites of the rickettsial insert are given in the recombinant plasmids. The corresponding restriction sites are aligned in the restriction map of the ε′ fragment at the top.

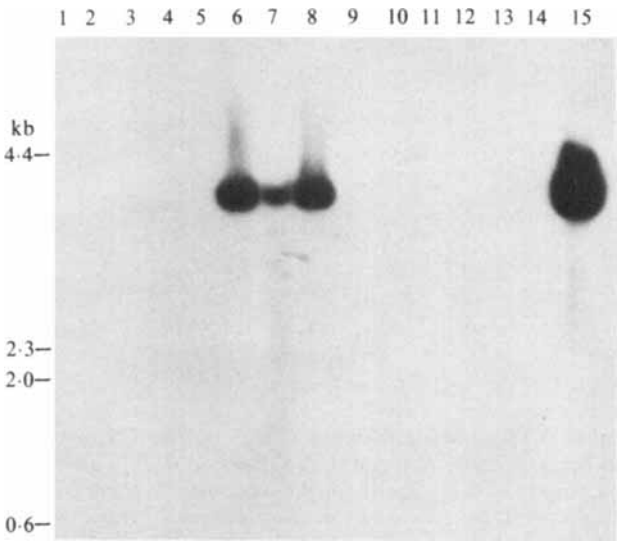

Fig. 2. Detection of the *cbbE′* gene by Southern blot hybridization of plasmid and chromosomal DNA isolated from typical isolates of the four known strains of *C. burnetii*. The 695 bp *Pst*I fragment of *cbbE′* served as the probe. All DNA samples were digested with *Eco*RI, except those in lanes 2, 3 and 8, which were digested with *Sma*I, *Sal*I, and *Eco*RI plus *Sal*I, respectively. Lanes: 1, pUC19; 2, pHK17; 3, pQH1; 4, QpH1; 5, Nine Mile chromosomal; 6, pQRS; 7, Priscilla chromosomal; 8, pQRS; 9, Dugway chromosomal; 10, QpDG; 11, Ko chromosomal; 12, S chromosomal; 13, G chromosomal; 14, L chromosomal; 15, pQME1.

Frazier, unpublished results). The 55 kDa protein was transcribed and translated from the insert regardless of its orientation in the vector, suggesting that transcription occurred from a rickettsial promoter, and that the promoter was recognized by the *E.coli* S30 lysate in the

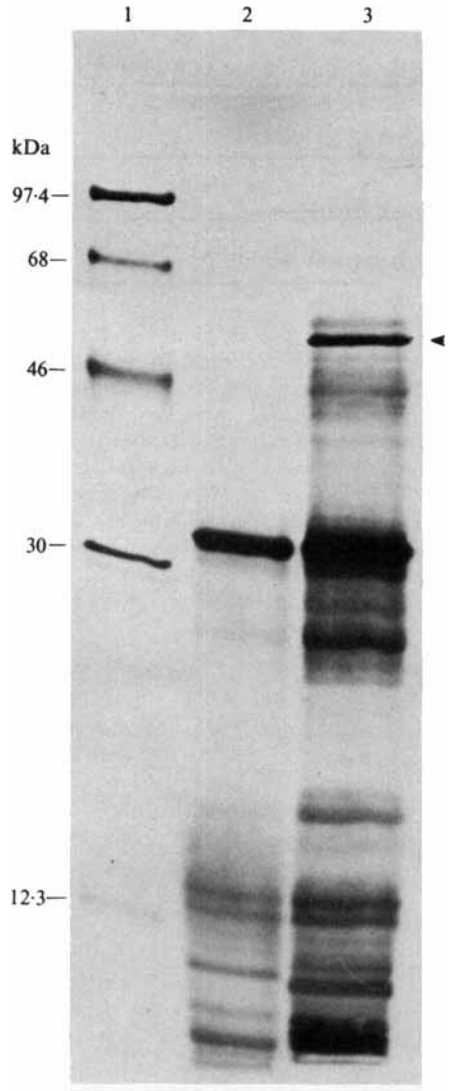

Fig. 3. IVTT of the *cbbE'* gene of *C. burnetii*. The [^{35}S]methionine-labelled proteins were separated by SDS-PAGE (12·5% acrylamide) and exposed to X-ray film overnight. The autoradiograph is shown. Lanes: 1, ^{14}C-labelled molecular mass standards; 2 and 3, IVTT products of the cloning vector pUC9 (lane 2) and of pQME3 (lane 3). The translated E′ protein is indicated by an arrowhead.

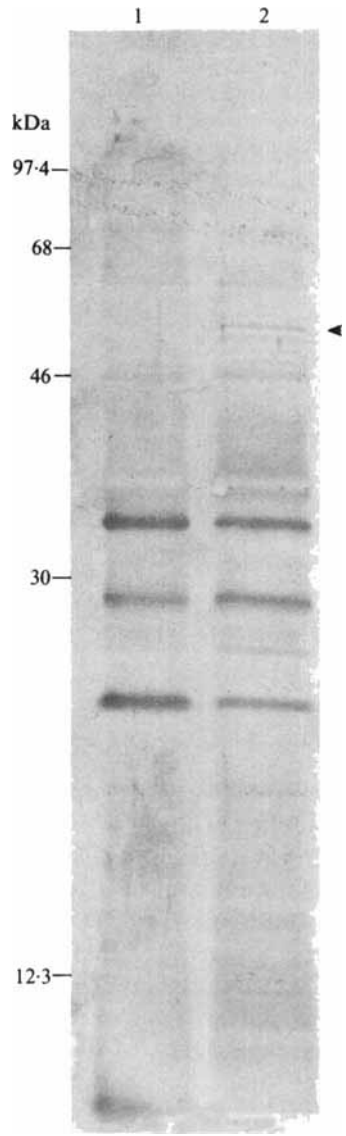

Fig. 4. *In vivo* expression of the *cbbE'* gene of *C. burnetii* in *E. coli*, and immunoblot detection of the expressed E′ protein with rabbit antiserum generated against purified OM of the Priscilla isolate of *C. burnetii*. Lanes: 1, *E. coli* SG932 containing the cloning vector pUC9; 2, *E. coli* SG932 containing pQME3. The E′ protein is indicated by an arrowhead.

IVTT system. The gene coding for the 55 kDa protein was designated *cbbE'*. *E. coli* SG932 containing pQME3 was capable of expressing the E′ protein *in vivo* (Fig. 4). The molecular mass (55 kDa) of the E′ protein expressed *in vivo* (Fig. 4) was the same as that of the E′ protein from IVTT analysis (Fig. 3). *E. coli* SG932 which contained only the cloning vector, pUC9, did not produce the 55 kDa protein (Fig. 4). In addition to the E′ protein, two cross-reactive bands of 51 and 53 kDa are uniquely present in *E. coli* containing pQME3 (Fig. 4, lane 2); these may be precursors or partial proteolysis products of the E′ protein. Three *E. coli* protein bands of approximately 25, 29, and 35 kDa were also recognized by the anti-OM antiserum (Fig. 4), and may represent protein species of *E. coli* which are antigenically related to OM protein(s) of *C. burnetii*.

Nucleotide sequence analysis of the cbbE' gene

To precisely locate the *cbbE'* ORF and determine its nucleotide sequence, a portion of the ε′ fragment was sequenced following the strategy shown in Fig. 5. An ORF long enough to code for the 55 kDa E′ protein

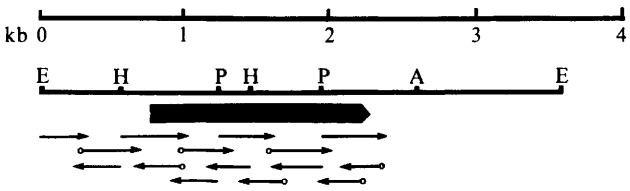

Fig. 5. DNA sequencing strategy. The region of cloned *C. burnetii* QpRS DNA which was sequenced is shown, and the sequencing strategy is indicated by arrows. Plain arrows denote subcloned segments primed with commercially available primers; arrows with open circles denote segments primed internally with oligonucleotides synthesized using the deduced sequence. A large arrow designates the location and direction of the ORF of *cbbE'*. Restriction sites are indicated: A, *Acc*I; E, *Eco*RI; H, *Hin*cII; P, *Pst*I.

begins with an AUG codon 196 bp downstream from the *Hin*cII site at map position 0·55 and ends 293 bp downstream from the *Pst*I site at map position 1·95 (Fig. 5). The sequence of the 1485 bp *cbbE'* ORF was confirmed by sequencing both strands of DNA.

As the nucleotide sequence shows (Fig. 6) this ORF begins with an AUG initiation codon (position 196) and ends with a UAA termination codon at position 1681, followed by a second in-frame stop codon UAG at position 1687. Beginning 2 bases upstream of the AUG initiation codon, the ORF of *cbbE'* is preceded by a polypurine-rich sequence of GGAGAGA, representing a potential Shine–Dalgarno (SD) ribosome-binding sequence. In addition, a predicted promoter sequence TTTAAT-N_{15}-TATAAT occurs between positions 157 and 183 (Fig. 6).

The *cbbE'* ORF is followed by a region of dyad symmetry that could contribute to rho-factor-independent termination (Fig. 6; nucleotides 1691 to 1713), as predicted by the TERMINATOR program (University of Wisconsin Genetics Computer Group; Devereux *et al.*, 1984). The G+C content of the sequenced coding region was 39 mol%, which compares favourably with the value of 43 mol% for *C. burnetii* total genomic G+C (Weiss, 1982). A distinct preference for U and A was observed in the first or third position of the codon (Table 1).

Table 1. *Codon usage for the cbbE' gene of Coxiella burnetii, and the predicted amino acid composition of the gene product*

	Residues					Residues			
Amino acid	No.	% of total	Codons	No.	Amino acid	No.	% of total	Codons	No.
Ala	35	7	GCU	12	Leu	56	11·5	UUA	15
			GCC	6				UUG	11
			GCA	11				CUU	10
			GCG	6				CUC	7
Arg	15	3	CGU	4				CUA	9
			CGC	2				CUG	4
			CGA	4	Lys	50	10	AAA	35
			CGG	0				AAG	15
			AGA	3	Met	3	1	AUG	3
			AGG	2	Phe	22	4·5	UUU	13
Asn	18	4	AAU	15				UUC	9
			AAC	3	Pro	22	4·5	CCU	8
Asp	18	4	GAU	13				CCC	5
			GAC	5				CCA	4
Cys	7	1·5	UGU	5				CCG	5
			UGC	2	Ser	36	7	UCU	8
Gln	22	4·5	CAA	12				UCC	8
			CAG	10				UCA	4
Glu	41	8	GAA	28				UCG	6
			GAG	13				AGU	6
Gly	26	5	GGU	6				AGC	4
			GGC	4	Thr	22	4·5	ACU	5
			GGA	13				ACC	5
			GGG	3				ACA	7
His	11	2	CAU	9				ACG	5
			CAC	2	Trp	4	1	UGG	4
Ile	41	8	AUU	16	Tyr	20	4	UAU	14
			AUC	10				UAC	6
			AUA	15	Val	26	5	GUU	7
								GUC	7
								GUA	7
								GUG	5

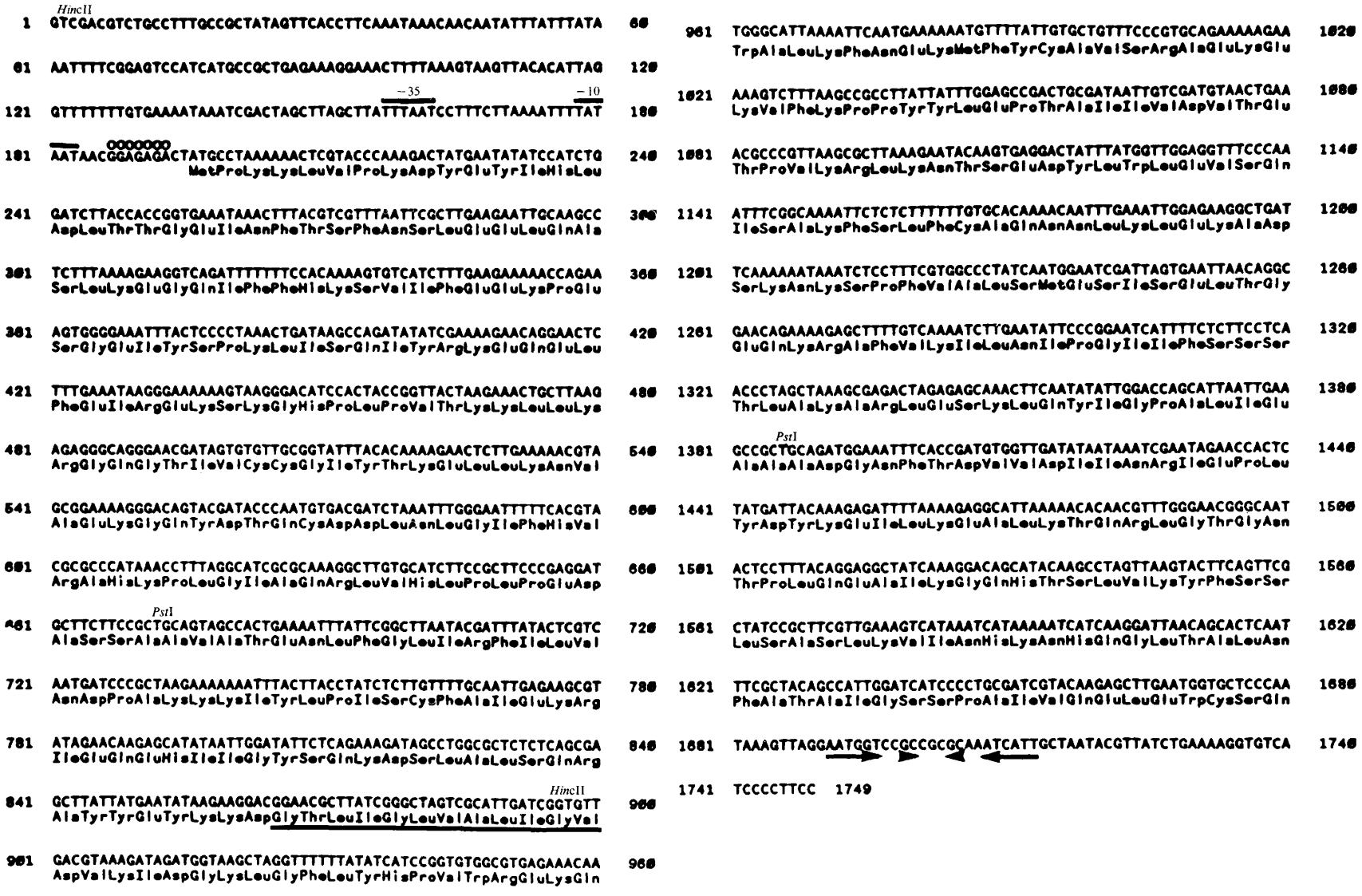

Fig. 6. Nucleotide sequence and deduced amino acid sequence (5′ to 3′) of the *C. burnetii cbbE′* gene. Nucleotide numbering begins with the *Hin*cII site located at map position 0·55 of the ε′ fragment (Fig. 1). Putative −35 and −10 promoter regions of *cbbE′* are overlined. A putative SD region is marked by open circles. The 3′ region of dyad symmetry is shown by arrows. The amino acid translation is shown below the sequence, beginning at the N-terminal methionine at nucleotide 196. The underlined area represents a 12-amino-acid region of hydrophobic character. Restriction sites are indicated.

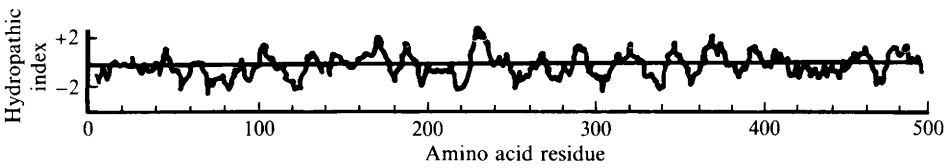

Fig. 7. Hydropathy plot of the *C. burnetii* E′ protein. Hydropathy was calculated by the method of Kyte & Doolittle (1982) with a span setting of nine amino acid residues. A hydrophobic segment is indicated when the plot extends into the upper half of the graph, while the lower half indicates a hydrophilic region of the protein.

Using the nucleotide sequence and the predicted amino acid sequence of the *cbbE′* gene, a computer search of the GenBank and NBRF data bases, respectively, was conducted; it revealed no DNA or amino acid homologies.

Features of the predicted amino acid sequence

The amino acid composition of the E′ protein as deduced from the nucleotide sequence is given in Fig. 6 and Table 1. The ORF of the *cbbE′* gene is capable of coding for a polypeptide of 495 amino acids with a predicted molecular mass of 55893 Da. This calculated molecular mass is in excellent agreement with the value of 55 kDa estimated for the E′ protein on SDS-PAGE. The isoelectric point of the highly charged, basic protein is predicted to be 8·7. One discrete hydrophobic domain from amino acids 224 to 235 is depicted in Fig. 6, and in the hydropathy plot (Fig. 7), indicating a possible transmembrane domain. A hydrophobic amino-terminal signal peptide similar to those seen in other transmembrane polypeptides (von Heijne, 1985) was not observed.

Discussion

To date, all *C. burnetii* isolates from chronic cases of Q fever contain the QpRS plasmid or have chromosomal DNA sequences with homology to QpRS. Isolates from acute cases contain the QpH1 plasmid. Virulence determinants that distinguish between chronic and acute strains of *C. burnetii* are unknown. We have used a molecular approach to determine whether DNA sequences that are unique to QpRS or QpH1 plasmids possibly code for virulence determinants that are important to the chronic or to the acute nature of the disease.

This report presents the first cloning, sequence, and expression, in *E. coli*, of a plasmid-encoded gene from *C. burnetii*. It is also significant that the *cbbE′* gene is unique to a plasmid harboured by chronic isolates associated with goat abortions and chronic endocarditis in humans (Samuel *et al.*, 1985). Whether the E′ protein serves as a virulence determinant in strains which contain the QpRS plasmid is unknown. The *cbbE′* gene of QpRS may code for a protein which has functional significance to those isolates which possess this plasmid type. Although a number of QpRS sequences have integrated into the chromosome of plasmidless-chronic isolates such as Ko, L, G or S (E. A. Savinelli & L. P. Mallavia, unpublished results), the *cbbE′* gene cannot be detected (Fig. 2), and appears to have been lost. The loss of *cbbE′* has no apparent effect on the ability of plasmidless-chronic isolates to cause chronic endocarditis, given the fact that all such isolates were obtained from infected heart tissues (Peacock *et al.*, 1983). One disease manifestation which is unique to those *C. burnetii* isolates which possess QpRS is their ability to cause abortion in goats and sheep (Samuel *et al.*, 1985). Whether the E′ protein is involved in causing such abortions is unknown.

The putative surface location of the E′ protein is supported by the fact that antiserum generated against *C. burnetii* OM can recognize the E′ protein on immunoblots. Further evidence is the observation that the E′ protein from IVTT comigrates on SDS-PAGE with a radio-iodinated surface protein which is unique to Priscilla but absent in the Nine Mile isolate (M. F. Minnick, R. A. Heinzen, M. E. Frazier & L. P. Mallavia, unpublished results). The radio-iodination data corroborate the DNA hybridization data which shows that *cbbE′* is present on QpRS from Priscilla, and is absent in the QpH1 plasmid or in the chromosomal DNA of the Nine Mile isolate (Fig. 2).

The predicted promoter regions for the *cbbE′* gene show considerable homology to the *E. coli* consensus promoter. The SD region is very similar to those found in *E. coli* (Gold *et al.*, 1981) and differs by an A-to-G transition in the first base of the sequence AGAGAGA from the SD region of the *Rickettsia rickettsii* 17 kDa antigen gene (Anderson *et al.*, 1987). The predicted −10 region or Pribnow box of *cbbE′* (TATAAT) precedes the AUG initiation codon by 12 bp, and is identical to the consensus sequence found in *E. coli* (Hawley & McClure, 1983). This −10 region of *cbbE′* differs from that of the 17 kDa antigen gene of *R. rickettsii* by a C-to-A

transition at base number 5 (Anderson *et al.*, 1987). The −35 region of *cbbE′* differs from the *E. coli* consensus sequence TTGACA (Hawley & McClure, 1983) by substitution of the last four bases with TAAT. The −35 region of *cbbE′* bears greater similarity to the −10 consensus sequence of *E. coli* and to the *R. rickettsii* −35 sequence (TTTACA) in a gene coding for a 17 kDa surface antigen (Anderson *et al.*, 1987).

The only previous report of sequenced DNA from *C. burnetii* is that of a heat-shock operon composed of *htpA* and *htpB* genes (Vodkin & Williams, 1988). Unfortunately, the −35 and −10 regions of the *htpA*/*htpB* operon comprise a heat-shock promoter which is not comparable to the *cbbE′* regulatory regions. The *htpB* gene does have a SD sequence that is similar to that of *cbbE′*, while *htpA* apparently lacks such a sequence.

The 15-base segment separating the −35 and −10 promoter sequences of *cbbE′* is marginal for recognition by *E. coli* RNA polymerase (Hawley & McClure, 1983), but it is similar to that of *R. rickettsii* (Anderson *et al.*, 1987). However, the ability of the IVTT system to recognize the *C. burnetii* promoter and ultimately produce the E′ protein from pUC19 recombinants containing the *ε′* insert in both orientations, one of which cannot utilize the *lacZ′* promoter (pQME2; Fig. 1), suggests that the *C. burnetii* promoter is recognized by the *E. coli* RNA polymerase. Detection of *in vivo* expression of the pQME3 construct in *E. coli* SG932 also suggests that the *C. burnetii* promoter is recognized by *E. coli* RNA polymerase.

In vivo expression of *cbbE′* was not detected in *E. coli* strains that were not Lon⁻, despite exhaustive attempts in several strains containing recombinant plasmids with *cbbE′*. Manipulation of the IPTG induction and/or harvest times following IPTG induction did not increase the yield of detectable amounts of the E′ protein. Although the E′ protein could be detected in *E. coli* SG932, the levels of *cbbE′* expression were disappointingly low, even on immunoblots (Fig. 4). This low level of detectable expression may be due to the suboptimal spacing between the −35 and −10 promoter regions of *cbbE′*. Analysis of E′ protein products by maxicell analysis (Sancar *et al.*, 1981) suggested that the synthesized *E′* protein was being rapidly degraded, despite a number of preventative measures (data not shown). Only the protease-deficient Lon⁻ *E. coli* strain SG932 provided an adequate host for accumulation of sufficient levels of intact E′ protein to be detected.

The U and A nucleotide preference in codons of *cbbE′* is similar to that of sequenced genes of other rickettsiae, including the citrate synthase gene of *Rickettsia prowazekii* (Wood *et al.*, 1987) and the *htpB* gene of *C. burnetii* (Vodkin & Williams, 1988). This result would be expected in an organism which has a 43 mol% G+C content (Weiss, 1982).

A discrete hydrophobic domain of 12 amino acids, flanked on both sides by an aspartic acid residue, was observed at approximately the centre of the E′ protein (between amino acid residues 224 and 235). This region may represent a membrane-spanning domain of the E′ protein, serving to anchor it to the OM. Twelve hydrophobic amino acids have been shown to be sufficient for a membrane-spanning domain (Adams & Rose, 1985). The aspartic acid residues occurring at both ends of the hydrophobic domain could possibly associate with the charged heads of the OM phospholipid bilayer (Adams & Rose, 1985). However, no apparent amino-terminal signal peptide with consensus to other prokaryotes (von Heijne, 1985) was observed.

The purpose of cloning the *cbbE′* gene was twofold. First, to obtain nucleotide sequence data for a *C. burnetii* plasmid gene, in order to compare its regulatory regions with those of genes of other rickettsiae and *E. coli*. Secondly, to characterize the gene and gene product because of their unique association with a plasmid, QpRS, harboured by chronic strains of *C. burnetii*. Although the presence of plasmid sequences appears to correlate with the virulence of *C. burnetii*, a specific gene coding for a virulence determinant has not yet been determined to be present on the plasmid. Characterization of strain-specific plasmid genes and proteins from *C. burnetii* strains with unique virulence potential is a logical first step to determining virulence function. The putative surface location of the E′ protein also suggests that it might eventually serve in detection tests. Finally, the unique DNA sequence of the *cbbE′* gene and the unique gene product it encodes could also serve as a diagnostic probe to distinguish between acute strains that contain QpH1 and chronic strains of *C. burnetii* that contain QpRS.

We would like to express our appreciation to Shirley Schmitt and Elizabeth Savinelli for their excellent technical assistance, and to Kathleen Postle for providing *E. coli* SG932. This work was supported by grant AI20190 from the National Institutes of Health (NIAID), and by the Northwest College and University Association for Science (University of Washington) under contract DE-AM06-76-RL02225 with the US Department of Energy.

References

ADAMS, G. A. & ROSE, J. K. (1985). Structural requirements of a membrane-spanning domain for protein anchoring and cell surface transport. *Cell* **41**, 1007–1015.

ANDERSON, B. E., REGNERY, R. L., CARLANE, G. M., TZIANABOS, T., MCDADE, J. E., FU, Z. Y. & BELLINI, W. J. (1987). Sequence analysis of the 17-kilodalton-antigen gene from *Rickettsia rickettsii*. *Journal of Bacteriology* **169**, 2385–2390.

BIRNBOIM, H. D. & DOLY, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Research* **7**, 1513–1523.

DAVIS, R. W., BOTSTEIN, D. & ROTH, J. R. (1980). *Advanced Bacterial Genetics*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

DEVEROUX, J., HAEBERILI, P. & SMITHIES, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Research* **12**, 387–395.

GOFF, S. A., CASSON, L. P. & GOLDBERG, A. L. (1984). Heat shock regulatory gene *htpR* influence rates of protein degradation and expression of the *lon* gene in *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America* **81**, 6647–6651.

GOLD, L., PRIBNOW, D., SCHNEIDER, T., SHINEDLING, S., SINGER, B. & STORMO, G. (1981). Translation initiation in prokaryotes. *Annual Review of Microbiology* **35**, 365–407.

HACKSTADT, T. (1986). Antigenic variation in phase I lipopolysaccharide of *Coxiella burnetii* isolates. *Infection and Immunity* **52**, 337–340.

HAWLEY, D. K. & MCCLURE, W. R. (1983). Compilation and analysis of *Escherichia coli* promoter DNA sequences. *Nucleic Acids Research* **11**, 2237–2255.

VON HEIJNE, G. (1985). Signal sequences the limits of variation. *Journal of Molecular Biology* **184**, 99–105.

HURLBERT, R. E. & GROSS, D. C. (1983). Isolation and partial characterization of the cell wall of *Pseudomonas syringae* pv. *syringae* HS191: comparison of outer membrane proteins of HS191 with those of two plasmidless derivatives. *Journal of General Microbiology* **129**, 2241–2250.

KIMBROUGH, R. C. III, ORMSBEE, R. A., PEACOCK, M., ROGERS, W. R., BENNETTS, R. W., ROOF, J., KRAUSE, A. & GARDNER, C. (1979). Q fever endocarditis in the United States. *Annals of Internal Medicine* **91**, 400–402.

KLEE, H. J., MILTON, G. P. & NESTER, E. W. (1983). Complementation analysis of *Agrobacterium tumefaciens* Ti plasmid mutations affecting oncogenicity. *Journal of Bacteriology* **150**, 327–331.

KYTE, J. & DOOLITTLE, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology* **157**, 105–132.

LAEMMLI, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature, London* **227**, 680–685.

MANIATIS, T., FRITSCH, E. F. & SAMBROOK, J. (1982). *Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

MOOS, A. & HACKSTADT, T. (1987). Comparative virulence of intra- and interstrain lipopolysacharide variants of *Coxiella burnetii* in the guinea pig model. *Infection and Immunity* **55**, 1144–1150.

PEACOCK, M. G., PHILIP, R. N., WILLIAMS, J. C. & FAULKNER, R. S. (1983). Serological evaluation of Q fever in humans: enhanced phase I titers of immunoglobulins G and A are diagnostic for Q fever endocarditis. *Infection and Immunity* **41**, 1089–1098.

SAMUEL, J. E., FRAZIER, M. E., KAHN, M. L., THOMASHOW, L. S. & MALLAVIA, L. P. (1983). Isolation and characterization of a plasmid from phase I *Coxiella burnetii*. *Infection and Immunity* **41**, 448–493.

SAMUEL, J. E., FRAZIER, M. E. & MALLAVIA, L. P. (1985). Correlation of plasmid type and disease caused by *Coxiella burnetii*. *Infection and Immunity* **49**, 775–779.

SANCAR, A., WHARTON, R. P., SELTZER, S., KACINSKI, B., CLARK, N. D. & RUPP, W. D. (1981). Identification of *uvrA* gene product. *Journal of Molecular Biology* **148**, 45–62.

SANGER, F., NICKLEN, S. & COULSON, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* **74**, 5463–5467.

SOUTHERN, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology* **98**, 503–517.

TOWBIN, H., STAEHELIN, T. & GORDON, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proceedings of the National Academy of Sciences of the United States of America* **76**, 4350–4354.

VIEIRA, J. & MESSING, J. (1983). The pUC plasmids, an M13 mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene* **19**, 259–268.

VODKIN, M. H. & WILLIAMS, J. C. (1988). A heat shock operon in *Coxiella burnetii* produces a major antigen homologous to a protein in both mycobacteria and *Escherichia coli*. *Journal of Bacteriology* **170**, 1227–1234.

VODKIN, M. H., WILLIAMS, J. C. & STEPHENSON, E. H. (1986). Genetic heterogeneity among isolates of *Coxiella burnetii*. *Journal of General Microbiology* **132**, 455–463.

WEISS, E. (1982). The biology of rickettsiae. *Annual Review of Microbiology* **36**, 345–370.

WOOD, D. O., WILLIAMSON, L. R., WINKLER, H. H. & KRAUSE, D. C. (1987). Nucleotide sequence of the *Rickettsia prowazekii* citrate synthase gene. *Journal of Bacteriology* **169**, 3564–3572.